Biological Inspiration for Multiple Memories Implementation and Cooperation

Frédéric Alexandre

Biological Inspiration for Multiple Memories Implementation and Cooperation

Frédéric Alexandre

LORIA/INRIA-Lorraine
BP 239
F-54506 Vandoeuvre, France

Abstract. Biological inspiration has led to the design of many connectionist models and mechanisms. Among them, memorization mechanisms are of particular importance to endow biologically inspired systems with efficient and consistent adaptive abilities. In this paper, we report recent modelling works of this kind, implementing procedural, episodic and working memories and making them cooperate for autonomous agent navigation.

Introduction

Classical connectionist modelling has explored a variety of adaptive mechanisms [8] and has allowed many impressive results in the domain of pattern matching, control, prediction, etc. Coming back to biological inspiration for the design of new models or mechanisms can be justified by at least three reasons. First, original and robust adaptative properties can be extracted from the observation of animal behavior. Second, designing adaptive mechanisms in relation to their neuronal substratum can give a framework to make these mechanisms cooperate in integrated tasks and not act alone on artificial problems. Third, the reference to animal behavior favors the implementation of ecological tasks that are too rarely investigated. In this paper, this kind of researches is illustrated through the design of three different kinds of memories (namely, procedural, episodic and working memories) and their interaction in a cerebral framework for endowing an autonomous agent with navigation and environment exploration capacities.

1 Procedural memory

Procedural learning is the ability to learn functions or procedures. This ability is reported as a property of the associative posterior cortex which can learn for example sensorimotor coordination. This learning is very slow and many trials are necessary for example to learn how to guide one's hand with one's eye. This kind of learning can be somehow related to the slow statistical learning performed by feed-forward networks like multi-layer perceptrons. Indeed, a pattern matching can be learned between an input space and an output space and can correspond to learn the transformations necessary to go from one reference frame (for example retinal) to another (for example muscular).

From this basic property, illustrated by classical connectionist models, biologically inspired models of the associative posterior cortex can bring additional structural and adaptive properties. Concerning structure, an important improvement can be brought to information representation with the implementation of cortically inspired areas [1]. Such areas, close to Kohonen self-organizing maps [10], implement batteries of filters and can accordingly be learned with Kohonen-like competitive mechanisms. Receiving information from sensors or actuators, such monomodal areas can build a mapping representing basic prototypes of perceptive or motor events, which is a first step in the internal representation of the external world.

Procedural learning will be performed in associative areas, combining monomodal or other associative maps. Basically, this learning lies on hebbian associative mechanisms, computing correlations between events in associated areas. Thus, for example, an association can be learned between the image of the arm, represented in retinal coordinates into a visual area and its sensation, represented in muscular coordinates into a sensory area [3]. Additional biologically inspired mechanisms have been proposed (see for example [4]) and can lead to such interesting properties as generalization, robustness to noise, transfer of learning and association of sequences of events. This latter property is very important, for example to learn the consequences of one's actions on the environment. Such a learning has been used to build a reactive architecture allowing an autonomous agent to move, in order to reach a goal it could see [4].

2 Episodic memory

In addition to this slow procedural learning, humans and animals are obviously able to memorize after just one or two experiences, episodes of events they have just lived. The hippocampus has been reported [14] as the neuronal structure able to perform episodic learning. Basically, this learning can be described as follows[12] : As explained above, the posterior cortex includes many associative areas representing multimodal information. All these associative areas project in the entorhinal cortex which thus gives a snapshot of the current state of the associative posterior cortex. Thanks to its recurrent connectivity, the CA3 structure in the hippocampus is then able to store a prototype of this activity. Similarly, Hopfield-like recurrent neural networks [9] are reported as able to store prototypes of activity as stable states and act as auto-associative memory to converge from a close initial state toward this stable state. In the same way, the hippocampus can detect that the current state of the associative posterior cortex is novel or that a close state has been recently experienced. In the former case, the state can be memorized in a distributed way in the hippocampal structure. In the latter case, the memorized event, including its spatial and temporal context, can be retrieved and sent back to the cortex [13]. Such a short term, rapid memorization can allow an autonomous agent to reach a goal that was seen before, even it is no longer viewed [7]. It was also tested, in cooperation with a model of posterior cortex, as a supervisor of this latter structure [11], thus allowing for an internal cooperation between episodic and procedural memories.

3 Working memory

Both episodic and procedural memories can reveal insufficient to guide an autonomous agent if obstacles or other problems prevent it from directly reaching a goal. In this case, a working memory is necessary to manipulate goals and subgoals and perform planning. The prefrontal cortex is known as the neuronal structure responsible for this kind of reasoning [6]. Basically, a mechanism has been proposed in [2] to explain how sensorimotor goals and subgoals, detected in the posterior cortex, can be controlled over time to organize behavior. Neurons in the prefrontal cortex are known to have lasting activities during such planning behavior. It is then proposed that they act like stacks, memorizing current and past states of the posterior cortex, together with the corresponding transitions. This knowledge can then be exploited to organize these states with regard to a goal and to trigger them when necessary. It was also proposed [6] that the same system could be used as a memory of the future and thus anticipate consequences of its actions. Such a model of prefrontal cortex, cooperating with a model of posterior cortex was implemented [5] to organize the behavior of an autonomous agent with regard to several motivations.

Discussion

The purpose of this paper was to propose a rapid overview of some recent biologically inspired models that exploit current knowledge about memory systems in the brain. We wish to underline here that these models are not only theoretical but that they were really implemented and used for the navigation of simulated or real robots. Even if it was not possible to gather all these data in this paper, the reader is invited to refer to the references to get more details about underlying equations or performances. Beyond these important data, the goal of this paper was to show that neurosciences accumulate more and more knowledge about brain functioning that can be more and more easily integrated in operational softwares. From this partial overview, it also appears that modelling the brain necessarily consists in designing cooperating modules, each performing a specific kind of information processing and memorization, as a result of phylogenesis. Thus, the cerebral framework prevents from only focusing on specific local mechanisms and underlines a distributed view of cooperating systems, from which new robust mechanisms can emerge. What we have proposed here is that wondering about memorization capabilities can be a fruitful way to build such an artificial cerebral system.

References

1. D. H. Ballard. Cortical connections and parallel processing: Structure and function. *The Behavioral and Brain Sciences*, 9:67–129, 1986.
2. Y. Burnod. *An adaptive neural network the cerebral cortex.* Masson, 1989.
3. R. Caminiti, P. Johnson, C. Galli, S. Ferraina, and Y. Burnod. Making arm movements within different parts of space: the premotor and motor cortical representation of a coordinate system for reaching to visual targets. In *J. Neurosci.*, volume 11, pages 1182–1197, 1991.
4. H. Frezza-Buet and F. Alexandre. Selection of action with a cortically-inspired model. In *Seventh European Workshop on Learning Robots*, pages 13–21, 1998.
5. H. Frezza-Buet and F. Alexandre. Modeling prefrontal functions for robot navigation. In *International Joint Conference on Artificial Neural Networks*, 1999.
6. J. M. Fuster. Frontal lobe and the cognitive foundation of behavioral action. In A.R. Damasio, H. Damasio, and Y. Christen, editors, *Neurobiology of Decision-Making.* Springer, 1996.
7. P. Gaussier, C. Joulain, J. P. Banquet, S. Lepêtre, and A. Revel. The visual homing problem: an example of robotics/biology cross fertilization. In *Robotics and Autonomous Systems*, 1999.
8. J. Hertz, A. Krogh, and R. Palmer. Introduction to the theory of neural computation. Addison Wesley, 1991.
9. J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. In *Proceedings of the National Academy of Sciences, USA*, pages 2554–2558, 1982.
10. T. Kohonen. *Self-Organization and Associative Memory.* Springer-Verlag, 1988.
11. J. Murre. *Hippocampus*, volume 6, chapter 6, TraceLink : A model of Amnesia and Consolidation of Memory, pages 675–684. David G. Amaral and Menno P. Witter, 1996.
12. R. O'Reilly and J. W. Rudy. Conjunctive representations in learning and memory : principles of cortical and hippocampal function. Technical report, Department of Psychology, University of Colorado, 1999.
13. N. Rougier and F. Alexandre. Spatial knowledge transfer between models of hippocampus and associative cortex. In *International Joint Conference on Neural Networtks.* IJCNN, July 1999.
14. L. R. Squire. Memory and the hippocampus: A synthesis from findings with rats, monkeys and humans. *Psychological Review*, 99(2):195–231, 1992.